Adam Rychtář; Štěpán Klapka; Lucie Kárná
Calculation of the detection properties in the binary symmetrical channel

In: Jan Chleboun and Pavel Kůs and Petr Přikryl and Miroslav Rozložník and Karel Segeth and Jakub Šístek (eds.): Programs and Algorithms of Numerical Mathematics, Proceedings of Seminar. Hejnice, June 21-26, 2020. Institute of Mathematics CAS, Prague, 2021. pp. 120–128.

Persistent URL: http://dml.cz/dmlcz/703107

Programs and Algorithms of Numerical Mathematics 20
J. Chleboun, P. Kůs, P. Přikryl, M. Rozložník, K. Segeth, J. Šístek (Eds.)
Institute of Mathematics of the Czech Academy of Sciences, Prague 2021

CALCULATION OF THE DETECTION PROPERTIES IN THE BINARY SYMMETRICAL CHANNEL

Adam Rychtář[1], Štěpán Klapka[1], Lucie Kárná[2]

[1] AŽD Praha, s.r.o
Žirovnická 2, Praha 10, Czech Republic
rychtar.adam@azd.cz, klapka.stepan@azd.cz
[2] Czech Technical University in Prague – Faculty of Transportation Sciences
Konviktská 20, Prague 1, Czech Republic
karna@fd.cvut.cz

Abstract: One of the important parts of railway signalling systems design is the safety of communication, achievable — among others — with the error detecting code. Getting evidence of quantitative safety targets, especially the probability of undetected error of the code, is a surprisingly complicated issue. We've analysed 2048 irreducible self-adjoint generator polynomials of the degree 32. More than 70 of these have a maximum probability of failure lower than the standard codes generally used. In this article we present the best of all codes we've analysed.

Keywords: error correcting codes, error probability, safety codes

MSC: 94B70, 94A40

1. Introduction

Evidence of the fulfilment of quantitative safety objectives is required when assessing the safety of railway signalling equipment, which is usually part of the documentation referred to as Safety Case. Quantitative targets are determined by the tolerable intensity of the dangerous failure (see EN 50129 [5] – Tolerable Functional (unsafe) Failure Rate TFFR).

Current European Standards ([5], [6]) recommend the use of formula (1) to estimate the probability of failure of a detection code:

$$p_{ud} = 2^{-c}, \tag{1}$$

where c denotes the number of *redundant* (or *control*) bits.

This however assumes that the code used is shown to have a property denoted as *proper* or *good* (for explanation see Paragraph 1.2 below). The examples given in

DOI: 10.21136/panm.2020.12

Chapter 2 show that these assumptions do not apply to commonly used detection codes (safety and transmission codes).

In the case of the second example below, for the Ethernet link layer, many articles have been published that have tried to suggest improvements to the detection code used. One interesting group of detection codes are codes generated using irreducible self-adjoint polynomials. An analysis of the detection properties was performed for all these codes with particular lengths. Chapter 3 gives the results for the best of them, but even they do not meet the requirements for formula (1) to be applied. In Chapter 4, the authors discuss the procedure for finding a detection code for which this formula could be used with sufficiently small deviation.

1.1. Linear and cyclic codes

A *linear binary* (n, k)*-code* $\mathbf{C}$ is defined as an arbitrary k-dimensional subspace of the n-dimensional linear space $(\mathbf{Z}_2)^n$. Binary vectors are traditionally called *words*; words from the code $\mathbf{C}$ are *codewords*. Any linear (n, k)–code $\mathbf{C}$ can be described by its *generator matrix* G of the dimension $k \times n$, whose rows are exactly the words forming some basis of the subspace $\mathbf{C}$. A matrix H of the dimension $(n - k) \times n$ is called *parity-check matrix* of the code $\mathbf{C}$, if it has the following property: A word of the length n is a codeword of the code $\mathbf{C}$ if and only if its product with matrix H is zero word (i.e. zero vector).

The *dual code* $\mathbf{C}^\perp$ of a linear (n, k)-code $\mathbf{C}$ is the linear $(n, n - k)$-code, defined as a set of all words of the length n being orthogonal to all codewords of the code $\mathbf{C}$, i.e. $\mathbf{C}^\perp = \{u \in (\mathbf{Z}_2)^n \mid \sum_{i=1}^{n} u_i v_i = 0 \ \forall v \in \mathbf{C}\}$. A generator matrix for the dual code $\mathbf{C}^\perp$ is a parity-check matrix for the original code $\mathbf{C}$ and vice versa.

The *cyclic code* is a linear code closed in respect to the circular shifts. That means, for every codeword $(a_0, a_1, \ldots, a_{n-1})$ the word $(a_{n-1}, a_0, a_1, \ldots, a_{n-2})$ is also a codeword. Codewords of a cyclic code of the length n can be written as formal polynomials $\mathbf{p}(x)$ of the degree lesser than n. Then multiplication by x corresponds to a cyclic shift. Every nontrivial (i.e. containing more than one word) cyclic (n, k)-code contains exactly one polynomial $\mathbf{g}(x)$ of the minimal possible degree among all non-zero polynomials, its degree being $n - k$. The polynomial $\mathbf{g}(x)$ is *generator polynomial* of the code $\mathbf{C}$. The dual code to the cyclic code is cyclic code as well.

Let $\mathbf{C}$ be a linear binary (n, k)-code and M be some m-elements subset of the set of indices $\{0, 1, 2, \ldots, n-1\}$. Define a set $\mathbf{C}(M)$ of all codewords that have zeros in all components from M: $\mathbf{C}(M) = \{u \in \mathbf{C} \mid u_i = 0 \, \forall i \in M\}$. By omitting the components from M we obtain a linear binary code $\mathbf{C}_{n-m}$ of the length $n - m$. The code $\mathbf{C}_{n-m}$ is called *shortened code* of the code $\mathbf{C}$.

If the code $\mathbf{C}$ is a cyclic code, its shortened codes are in practice referred to as *shortened cyclic codes*, although they are almost never cyclic.

More details concerning cyclic codes and their construction can be found in Berlekamp's book [1] for example.

1.2. Detection properties

The basic measure for detection ability of the code is its minimal distance. For linear binary codes the *minimal (Hamming's) distance* is defined as minimal weight of nonzero codeword, where *weight* of the word is the number of its nonzero symbol.

It is a well known fact that undetected errors of the given linear (n, k)-code are all its nonzero codewords. Consequently, the linear (n, k)-code with minimal distance equal to d can detect all errors up to multiplicity $d-1$ (i.e. up to $d-1$ wrong bits).

Of course, this code can detect some of the errors of higher multiplicity, and at this point there are big differences amongst codes with a certain minimal distance. For a more detailed approach the *binary symmetrical channel* model is used.

A binary symmetrical channel (BSC) is a simple probability model based on an independent transmission of single bits (binary symbols). The probability p_e that the bit changes its value during the transmission (*bit error rate*) is the same for both possibilities (from 0 to 1 and vice versa).

The following formula (2) for the probability of undetected error p_{ud} is valid for a linear (n, k)-code in the binary symmetrical channel:

$$p_{ud}(p_e, n, A^n) = \sum_{i=1}^{n} A_i^n p_e^i \left(1 - p_e\right)^{n-i} \tag{2}$$

where A_i^n is the number of codewords of the weight i. The vector A^n is called *weight distribution* of the code.

The key problem is what maximum value the function $p_{ud}(p_e, n, A^n)$ can take. Setting the value p_e to one half, the formula (2) gives the following local upper bound, valid on some neighbourhood of $1/2$:

$$p_{ud}\left(\frac{1}{2}, n, A^n\right) < 2^{-(n-k)}.$$

This upper bound is independent of the code weight distribution and even of the codewords length n; it depends only on the number of *redundant bits* $n-k$.

Unfortunately, our intuitive expectation that with the decrease of bit error rate p_e the probability of undetected error p_{ud} decreases as well, is not always valid. In many cases the value $p_{ud}\left(1/2, n, A^n\right)$ is only a local maximum of the p_{ud} function on some neighbourhood of the value $1/2$.

The (n, k)-code is said to be *proper* if its function $p_{ud}(p_e, n, A^n)$ is monotone for p_e from the interval $\langle 0, 1/2\rangle$. However, for the hazard rate calculation it is necessary to know an upper bound of the p_{ud} only. The monotonicity of this function is not crucial for this purpose. The (n, k)-code is called *good* if the value $2^{-(n-k)}$ is the upper bound of the function $p_{ud}(p_e, n, A^n)$ on the whole interval $\langle 0, 1/2\rangle$.

1.3. Weight enumerator

To get more precise information about probability p_{ud} we focus on calculation of the code weight distribution A_n. The *weight enumerator* $\mathbf{pw}(x, n, A^n)$ is essential for

these calculations. It describes the distribution of codewords weights by the following formal polynomial:

$$\mathbf{pw}(x, n, A^n) = \sum_{i=0}^{n} A_i^n x^i. \tag{3}$$

To get a more effective calculation, it is useful to implement the MacWilliams Identity, which links the weight enumerators of the given code $\mathbf{pw}(x, n, A^n)$ and of its dual code $\mathbf{pw}(x, n, B^n)$. The following formula (4) is the form of MacWilliams Identity for binary codes:

$$2^k \mathbf{pw}(x, n, B^n) = (1+x)^n \mathbf{pw}\left(\frac{1-x}{1+x}, n, A^n\right). \tag{4}$$

The practical advantage of this procedure is that the dual code has far less codewords ($2^{n-k} \ll 2^k$), and thus the computation of its weight distribution is significantly faster.

2. Examples

Following graphs demonstrate properties of the BSC model for some linear binary codes. What we call a code in technical practice, in fact is a family of shortened cyclic codes of a single cyclic code, identified by its generator polynomial.

As a result, the following graphs are three-dimensional: the X-axis represents length of the code n, the Y-axis represents bit error rate p_e, and the (vertical) Z-axis represents the respective probability of undetected error $p_{ud}(p_e, n, A^n)$. Note that scales on the Y and Z axes are different.

2.1. Example 1: ETCS code

Because our research was motivated by the development of the railway signalling systems, the first example is from this application area. The discussed code is used for safety communication in the European Train Control System (ETCS). The code was designed using BCH code construction (see e.g. [1]) with minimal distance equal to 6 and construction length of 32767 bits. Its generator polynomial of the code is $x^{32} + x^{30} + x^{27} + x^{25} + x^{22} + x^{20} + x^{13} + x^{12} + x^{11} + x^{10} + x^8 + x^7 + x^6 + x^5 + x^4 + 1$.

The behavior of the function $p_{ud}(p_e, n, A^n)$ for the ETCS code is shown by the graph in Fig. 1. Note that the value of the local upper bound at $p_e = 1/2$ is $2^{-32} \approx 2.3 \cdot 10^{-10}$, which is (in the scale used in the graph) near to zero. It is obvious that the ETCS code is neither good nor proper.

2.2. Example 2: Ethernet

The other heavily used code is transmission code of the Ethernet link layer (IEEE 802.3). It is a shortened cyclic code with generator polynomial $x^{32} + x^{26} + x^{23} + x^{22} + x^{16} + x^{12} + x^{11} + x^{10} + x^8 + x^7 + x^5 + x^4 + x^2 + x + 1$. This is a primitive polynomial and generates a cyclic Hamming code (for explanation see [3]). Hamming

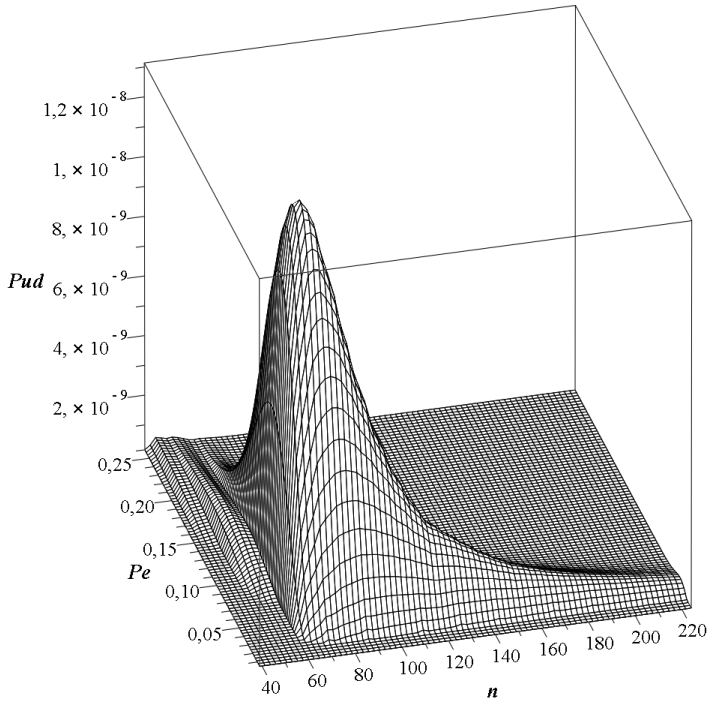

Figure 1: Probability of undetected error for the ETCS code. (X-axis: length of the code n, Y-axis: bit error rate p_e, vertical axis: probability of undetected error p_{ud}.)

codes are proper, but shortened Hamming codes are not. Minimal distance of this code drops to 5 for codeword length 269 bits already. Detection properties of this code are shown by the graph in Fig. 2. The maximal values of the p_{ud} are lower than those of the ETCS code, and they are only about three times higher than 2^{-32}. However, the code is not good nor proper.

3. Examined codes and results of calculations

Transmission codes with 32 bit redundancy are further discussed by Koopman in his work [4]. His goal was to find transmission codes that maximise the minimal distance for the largest possible code lengths. Two of discussed codes were generated by irreducible self-adjoint polynomials. Shortened cyclic codes generated by this group of non-primitive polynomials have a minimal distance equal to five up to a codeword length of 65536 bits. This property allows correction of up to two bit errors for the lengths mentioned. However, the maximum values of the p_{ud} for the codes mentioned by Koopman are several times higher than for the codes referred in Sections 2.1 and 2.2.

Another example of code generated by irreducible self-adjoint polynomial was published by Castagnoli in [2]. The code proposed in his article is — regarding

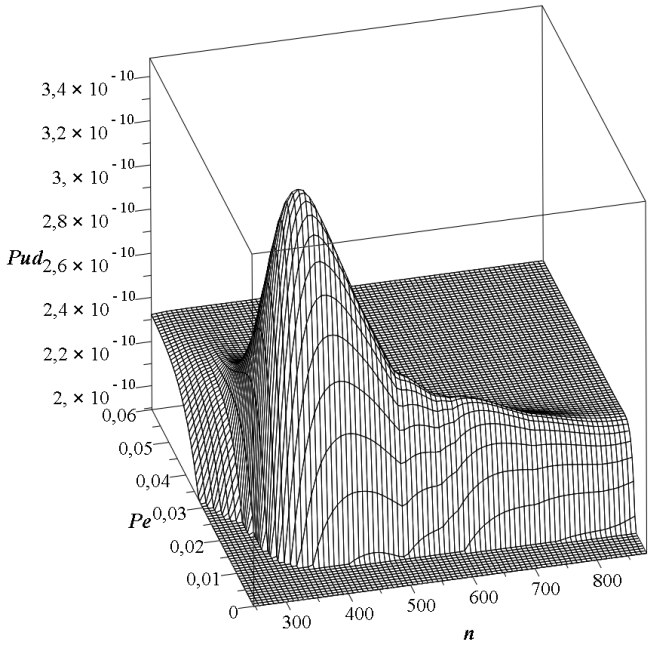

Figure 2: Probability of undetected error for the Ethernet link layer code. (X-axis: length of the code n, Y-axis: bit error rate p_e, vertical axis: probability of undetected error p_{ud}.)

the p_{ud} function — a better choice than the polynomials proposed by Koopman, but it does not reach the quality of the polynomial for Ethernet.

We were looking for codes with big minimal distance for long codewords, and with best possible probabilistic results in the BSC model. Inspired by the mentioned studies [4] and [2], we analysed 2048 irreducible self-adjoint generator polynomials of the 32th degree. Considered codeword lengths were multiples of 8 from 40 to 65536.

Performed calculations pointed out that not one of these polynomials generate proper code for every considered codeword length. However, many of them have a lower maximum of the p_{ud} function than the Ethernet link layer code. The best of them (denoted as c1798) is generated by the polynomial $x^{32}+x^{31}+x^{30}+x^{29}+x^{23}+x^{20}+x^{19}+x^{17}+x^{16}+x^{15}+x^{13}+x^{12}+x^9+x^3+x^2+x+1$. Its detection properties are shown by the graph in Fig. 3. The plateau of the graph is near value 2^{-32}. Notice two local maxima of the p_{ud} function.

Table Tab. 1 summarises more detailed information about properness of the c1798 code. In the first column there are intervals of codeword lengths, for which the shortened c1798 code is/isn't proper. The fourth column indicates codeword length (in this interval) with the highest value of the maximum of the p_{ud} function. The ratio between this value $\max(p_{ud})$ and the value of $p_{ud}(1/2)$ is given in the third column of the table.

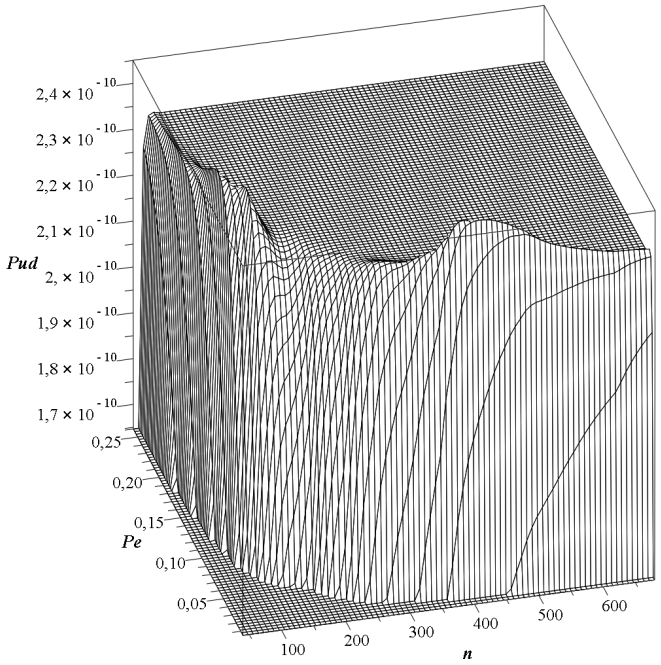

Figure 3: Probability of undetected error for the code c1798. (X-axis: length of the code n, Y-axis: bit error rate p_e, vertical axis: probability of undetected error p_{ud}.)

As the case turns out to be so far, the properness of shortened cyclic codes for all required lengths is a very rare feature. It has been found from previous analyses of 8th and 16th degree primitive polynomials, that there are some 8th degree polynomials with this property, but not a single 16th degree polynomial. Nevertheless, these polynomials (cyclic Hamming codes) are most often used in current error correcting codes in SRAM memories.

After all, the c1798 code is not too far from the proper code. For comparison, the following figure Fig. 4 demonstrates the p_{ud} function of the proper code (graph on the left) and of the c1798 code (graph on the right) with the same scale on the vertical axis.

4. Further research

We will focus on the extensive search of suitable generator polynomials among primitive polynomials of the 32th degree in the nearest future. Since there are 67108864 primitive polynomials of the 32th degree, first there must be procedures designed for reducing the number of examined polynomials. In addition, it is necessary to make maximum use of the current possibilities of calculations parallelisation.

In the long run, we want to improve techniques being used so that the area of generator polynomials of the 48th degree is accessible for computing. The current obstacle is the long computational time of the weight distribution, and huge amount of explored polynomials.

Interval (n)	Proper code	$\max(p_{ud})/p_{ud}(1/2)$	Worst length
40	NO	1.002039674108	40
48 - 64	YES	1	—
72 - 160	NO	1.042927066522	112
168 - 232	YES	1	—
240 - 13000	NO	1.054302693462	408
13008 - 65536	YES	1	—

Table 1: Lengths of codewords, where the code c1798 is proper.

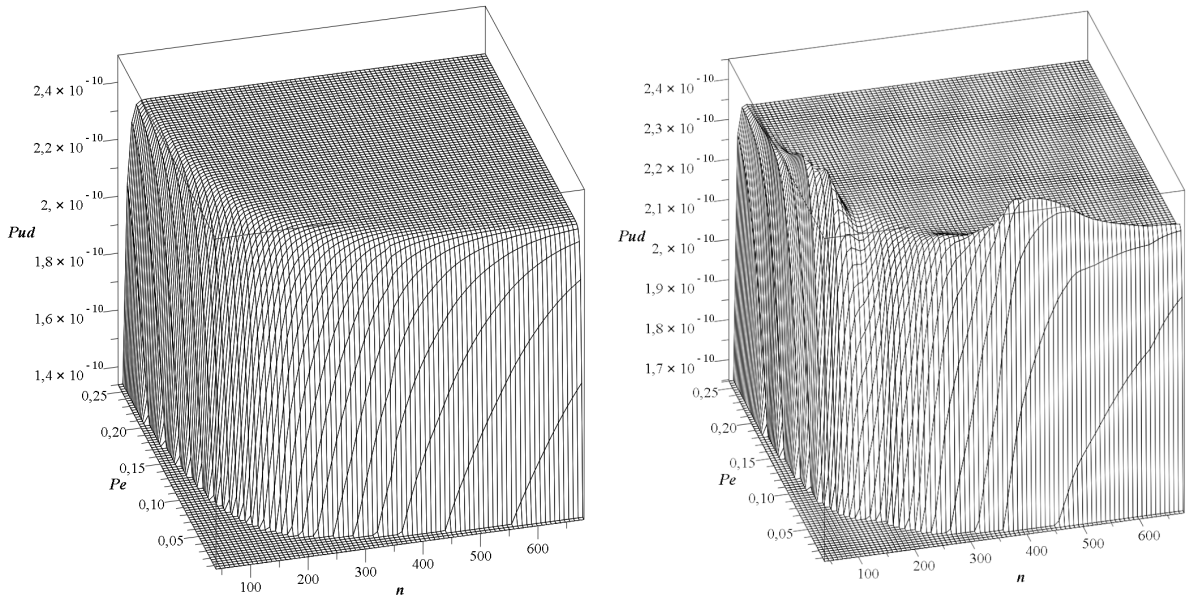

Figure 4: Comparison of the probability of undetected error for the proper code (on the left) and for the code c1798 (on the right).
(X-axis: length of the code n, Y-axis: bit error rate p_e, vertical axis: probability of undetected error p_{ud}.)

5. Conclusion

Probabilities of undetectable errors of cyclic codes with generator polynomials of the 32nd degree were calculated using a BSC model. A set of codes with irreducible and self-adjoint generator polynomials have been investigated.

These codes have a large minimal distance for codeword lengths up to 65536 bits. None of them are good nor proper for the entire range of lengths analysed, but some of them have a maximum probability of failure lesser than the standard codes used.

References

[1] Berlekamp, E. R.: *Algebraic coding theory.* McGraw-Hill, New York, 1968.

[2] Castagnoli, G., Braeuer, S. Herrman, M.: Optimization of cyclic redundancy-check codes with 24 and 32 parity bits. *IEEE Trans. Comm.* **41** (June 1993).

[3] Hankerson, D. R. et al.: *Coding theory and cryptography.* Marcel Dekker, 2000.

[4] Koopman, P.: 32-Bit Cyclic Redundancy Codes for Internet Applications. In: *Proc. Int. Conf. on Dependable Sys. and Networks*, pp. 459–468, doi: 10.1109/DSN.2002.1028931. Washington, 2002.

[5] Railway applications — Communication, Signalling and Processing Systems — Safety related Electronic Systems for Signalling. EN 50129, CENELEC, 2019.

[6] Railway Applications — Communication, Signalling and Processing Systems — Safety-related Communication in Transmission Systems. EN 50159, CENELEC, 2011.